



June 13, 2025

Hon. Ona T. Wang
Daniel Patrick Moynihan U.S. Courthouse
500 Pearl Street
New York, NY 10007-1312

311 N. Aberdeen Street
Chicago, Illinois 60607
(312) 243-5900
www.loevy.com

RE: *In re: OpenAI, Inc., Copyright Infringement Litigation, 1:25-md-03143-SHS-OTW; this document relates to The Center for Investigative Reporting, Inc. v. OpenAI, Inc. et al., Case No. 1:24-cv-04872-SHS-OTW*

Dear Judge Wang:

During the May 27, 2025 conference, the Court stated that if OpenAI files a renewed motion to compel additional custodians, the Court “need[s] to see *specific* allegations, *specific* documents, statements, admissions that indicate that these custodians are non-cumulative, non-duplicative responsive documents, and that it’s proportional to the needs of the case.” (See Ex. 1 at 156:9-15 (emphasis added).) OpenAI has failed to provide this material. In its renewed motion, OpenAI has plainly regurgitated the same exact arguments raised in its original motion. (ECF 250.) It has also failed to identify any specific statements or documents supporting its position to add new ESI custodians it seeks.¹

As for its faulty reasoning, OpenAI contends that CIR is manually collecting documents “as a substitute for the identification of custodians and search terms ordered by the ESI Order.” (ECF 119 at 1.) Once again, this is incorrect. Rather, CIR has agreed to run 17 broad search terms (terms that were proposed by OpenAI and agreed to by CIR, without any modification) across each agreed-upon custodian. (Ex. 2.) CIR’s manual collection of documents does not serve as a “substitute” for its custodial searches. Instead, the two methods of collection are done in conjunction with one another. *Thomas v. City of New York*, cited by OpenAI in its motion (ECF 119 at 1), is inapposite to CIR’s collection efforts, as defendants’ counsel in *Thomas* sought to exclude all “GroupMe chats between various parties” based solely on “their clients’ statements that these chats do not contain pertinent or relevant information.” 336 F.R.D. 1, 5 (E.D.N.Y. 2020). As CIR has explained, some categories of records are retained in discrete files and folders (e.g., relevant employment and work-for-hire agreements), and CIR has manually collected from those locations. To the extent OpenAI identifies anything that appears to be missing (for example, should there be authors of asserted works for which CIR has not produced an employment agreement), CIR will investigate and take reasonable steps to remedy, including targeted ESI searches if appropriate. But OpenAI has not done so.

¹ As background, CIR has designated 3 individual document custodians (the CEO, CFO, and VP of Media Sales/Business Development) who are the key decisionmakers at CIR regarding any issues relevant to this case, as well as 4 other custodial accounts (permissions@motherjones.com, registrations@motherjones.com, copyright@motherjones.com, and republish@revealnews.org). These individual custodians represent more than 2.5% of CIR’s employees, whereas OpenAI has designated only 0.5% of its workforce as custodians. (See ECF No. 253.)



June 13, 2025

Page 2 of 4

While OpenAI has plainly not fulfilled any of the Court's requirements for adding custodians, CIR reiterates again why these particular custodians are duplicative of other custodians and collection efforts:

Robert Wise (Online Technology Director). OpenAI contends that Wise should be added as a custodian because he may possess information relating to CIR's use of exclusion protocols and the display of CMI on its websites, as well as historical versions of and modifications to CIR's robots.txt files. (ECF 119 at 1-2.) As CIR's custodial 30(b)(6) deposition witness and current custodian, Monika Bauerlein (CIR's CEO), testified, Wise is a technical employee "responsible for the teams that maintain [CIR's] websites." (ECF 250-5 at 20:12-17.) Wise is not responsible for setting or changing any policies with respect to CIR's use of exclusion protocols or adding CMI to CIR's works. Rather, Bauerlein would be included on such communications to the extent they "affect[] any significant changes in the user experience or in [CIR's] organizational policies." (*Id.* at 184:13-20.) In addition to running broad search terms across its current custodians related to these issues, including, for example, *((webmaster* OR crawl* OR *robot* OR "robots.txt" OR exclude* OR exclus*) AND (scrap* OR paywall OR cach* OR access* OR *index))* (Ex. 2), CIR has agreed to manually collect responsive materials including, for example, CIR's robots.txt files. OpenAI has failed to identify any specific evidence that supports adding Wise as a custodian. To the extent that Wise possesses information relevant to this litigation, such information is (at best) duplicative.

Marla Jones-Newman (VP of People and Culture). OpenAI argues that Jones-Newman possesses documents responsive to "[CIR]'s ownership of the asserted works, including employment and work-for-hire agreements." (ECF 119 at 2.) OpenAI seeks to add Jones-Newman as a custodian because she has the "ultimate authority when it comes to [] employment agreements and keeping track of where they're stored and kept." (*Id.* (quoting ECF 250-5, Bauerlein Dep. Tr. at 179:3-8).) However, because CIR's relevant employment and work-for-hire agreements are stored in a central repository, CIR has agreed to manually search for and produce the agreements. There is no reason to add Jones-Newman as a custodian simply because she "keeps track" of where these agreements are "stored and located," when such agreements can be simply collected manually. Accordingly, OpenAI's request to add Jones-Newman as a custodian is unnecessary and not proportionate. If OpenAI believes that there are relevant agreements missing from CIR's productions, then OpenAI can identify them, and CIR will manually search for and produce such agreements, to the extent they exist, or will consider targeted custodial ESI searches if necessary.

Emily Harris (Director of Finance). OpenAI admits that "CIR has produced high-level analyses relating to advertising revenue and forecasts," but seeks to add Harris as a custodian because CIR's production is allegedly "largely missing data underlying those analyses, analyses of the financial value of the content, or analyses of other types of revenue streams, such as licensing revenue." (ECF 119 at 2.) However, Bauerlein (CIR's CEO, a current custodian) testified that Harris is "not typically" involved in creating these analyses in the first place. (ECF 250-5 at 160:21 – 161:2.) What's more, when asked "only once this analysis was created, would then, for example, Ms. Buckingham [CIR's CFO and existing custodian] be looped in; is that accurate?," Bauerlein answered, "***No. She and I would also have access to the underlying [financial] data and might look at it.***" (*Id.* at 161:23 – 162:2 (emphasis added).) Accordingly, any relevant communications within this category of documents runs through existing custodians (Bauerlein and Buckingham).

June 13, 2025

Page 3 of 4

OpenAI makes no reasonable attempt to explain why these custodians are insufficient. OpenAI also contends that “less than 20% of the documents CIR has produced relating to financial analyses are for Reveal.” (ECF 119 at 2.) Yet, Bauerlein testified that these analyses are not “regularly” performed for Reveal news articles because Reveal is “a much smaller website, and so there’s much more data to work with on the Mother Jones side.” (Ex. 3, Bauerlein Tr. at 162:21 – 163:3.) OpenAI has failed to identify any specific evidence to support the addition of Harris as a custodian.

Ruth Murai (Research Director). OpenAI seeks to add Murai as a custodian for issues relating to CIR’s fact-checking process for plagiarism and the use of AI. (ECF 119 at 3.) Bauerlein, however, testified that she (not Murai) is the decisionmaker regarding these topics. (See ECF 250-5 at 144-145 (testifying that she “would be in the loop on policies and changes to policies” regarding CIR’s fact-checking process, plagiarism detection efforts, and use of generative AI).) Indeed, CIR has confirmed based on discussions with Murai and Bauerlein that if Murai identified as part of the fact-checking process any incident of plagiarism or the unauthorized use of AI, Bauerlein would be promptly informed of the issue. Accordingly, whether Bauerlein is “involved in the ‘day-to-day’ fact-checking process for plagiarism or the use of AI,” as raised by OpenAI (ECF 119 at 3), is of no consequence because (i) Bauerlein oversees these issues (and thus possesses any relevant information), (ii) CIR has agreed to produce non-custodial documents reflecting this information, and (iii) CIR has run broad searches across its existing custodial accounts related to these issues (*see, e.g., Ex. 2 (searching the term *infring** OR *plagiar**)*). The fact that CIR has allegedly “produced no responsive documents apart from its employee handbooks” (ECF 119 at 3) is irrelevant. And, of course, OpenAI overlooks the possibility that no records evidencing plagiarism have been produced because CIR’s reporters have not engaged in plagiarism. OpenAI’s request to add Murai is unreasonable and not proportionate.

Michael Mechanic (Senior Editor). OpenAI seeks to add Mechanic as a custodian because he allegedly possesses documents relating to the use of third-party texts and sources, requests and licenses to third parties, and content-sharing licensing programs. (ECF 119 at 3.) CIR, however, has already agreed to manually collect responsive documents, and Bauerlein testified that one or more of CIR’s current custodians “would be on [] emails [relating to CIR’s use of third-party content] when it comes to agreeing with that publisher in the first place that they are wanting to be a part of this network.” (ECF 250-5 at 141:11-14.) Accordingly, to the extent CIR’s manual collection does not sufficiently capture responsive documents, such additional materials would be found in the existing custodial emails, as CIR’s current custodians are actively involved in these topics. Indeed, CIR has run broad searches across its existing custodial accounts related to these issues (*see, e.g., Ex. 2 (searching the terms ((contract* OR agreement* OR licens* OR arrangement) w/15 (reproduc* OR display OR derivative OR distrib* OR copy OR copie* OR syndicat* OR perform* OR prepar* OR train* OR excerpt* OR develop* OR summar* OR output* OR quot* OR refer* OR memoriz* OR generat* OR publish* OR work* OR artic* OR public* OR adopt* OR access* OR shar*)) (content OR contract* OR agreement* OR licens*) w/15 (subscription OR "news source" OR database* OR librar*)*). The fact that CIR has allegedly “produced documents describing only a handful of instances where CIR made a request for [third-party] use” (ECF 119 at 3) does not support OpenAI’s claim that Mechanic should be a custodian.

For the foregoing reasons, the Court should deny OpenAI’s renewed motion to compel CIR’s addition of the above custodians.

June 13, 2025

Page 4 of 4

Respectfully,

LOEVY & LOEVY

/s/ Kyle Wallenberg
Kyle Wallenberg

cc: All Counsel of Record (via ECF)